

Distributed Learning Policies for Power Allocation in Multiple Access Channels

Panayotis Mertikopoulos, Elena V. Belmega, *Member, IEEE*, Aris L. Moustakas, *Senior Member, IEEE*, and Samson Lasaulce, *Member, IEEE*

Abstract—We analyze the power allocation problem for orthogonal multiple access channels by means of a non-cooperative potential game in which each user distributes his power over the channels available to him. When the channels are static, we show that this game possesses a unique equilibrium; moreover, if the network's users follow a distributed learning scheme based on the replicator dynamics of evolutionary game theory, then they converge to equilibrium exponentially fast. On the other hand, if the channels fluctuate stochastically over time, the associated game still admits a unique equilibrium, but the learning process is not deterministic; just the same, by employing the theory of stochastic approximation, we find that users still converge to equilibrium.

Our theoretical analysis hinges on a novel result which is of independent interest: in finite-player games which admit a (possibly nonlinear) convex potential, the replicator dynamics converge to an ε -neighborhood of an equilibrium in time $\mathcal{O}(\log(1/\varepsilon))$.

Index Terms—Nash equilibrium; potential games; parallel multiple access channel; power allocation; replicator dynamics.

I. INTRODUCTION

IN view of the decentralized nature of future and emerging wireless networks, non-cooperative game theory has become an important tool to analyze distributed problems in networks whose nodes cannot be assumed to adhere to centrally controlled protocols. The main goal has been to develop policies and algorithms that nodes can use to optimize their resources (power, bandwidth, etc.) on their own, so, following [1], the questions that arise are *a*) whether there exist “equilibril” policies which are stable against unilateral deviations; *b*) whether these (Nash) equilibria are unique; and *c*) whether they can be reached by distributed learning algorithms that require only local information.

This work was supported in part by the Greek GSRT “Kapodistrias” project No. 70/3/8831, the GIS “Sciences de la Décision” (École Polytechnique–ENSAE–HEC), and the Pôle de Recherche en Economie et Gestion (UMR 7176). The work was initiated when E. V. Belmega was visiting the University of Athens during the summer of 2010 with the support of the L’Oréal program “For young women doctoral candidates in science 2009”. Part of this work was presented in the 5th International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS ’11), Paris, May 2011.

P. Mertikopoulos is with the Centre National de la Recherche Scientifique (CNRS) and the Laboratoire d’Informatique de Grenoble, France; at the time of submission, he was with the Economics Department of École Polytechnique, Paris, France (e-mail: panayotis.mertikopoulos@imag.fr).

E.V. Belmega is with ETIS/ENSEA – Université de Cergy-Pontoise – CNRS, Cergy-Pontoise, France; at submission, she was with the Alcatel-Lucent Chair on Flexible Radio in Supélec, France (email: belmega@ensea.fr).

A. L. Moustakas is with the Physics Department of the University of Athens, Greece (e-mail: arislm@phys.uoa.gr).

S. Lasaulce is with the Laboratoire des Signaux et Systèmes, Supélec, Gif-sur-Yvette, France (e-mail: samson.lasaulce@lss.supelec.fr).

Accordingly, an important paradigm which has attracted significant interest in the wireless communications literature concerns the allocation of power over orthogonal communication channels [1], [2], [3]. From a centralized viewpoint, this is a relatively well-studied subject, especially with respect to optimal power allocation schemes which allow users to reach the boundary of the rate region assuming full channel knowledge and central control [4], [5]. On the other hand, more recent examinations [6], [7], [8] focus on socially stable power allocation policies because, even if the globally optimal, capacity-achieving power profile is known and used, it might be unstable under deviations by selfish users (and thus useless in a decentralized setting).

In this paper, we consider the problem of uplink communication in multi-user networks consisting of several receivers that operate on distinct, non-interfering channels, and we focus on giving definitive answers to points (a)–(c) above, analyzing the equilibrial structure of the problem and its convergence aspects. Despite its apparent simplicity, this parallel multiple access channel (PMAC) model has several relevant applications such as, for instance, in 802.11-based wireless local area networks (WLANs) with non-overlapping channels [9], [10], distributed soft handoffs in cellular systems [11], distributed power allocation in digital subscriber lines (DSLs) [12], and, finally, in throughput-maximizing power control in multi-carrier code division multiple access (MC-CDMA) systems [13].

Our analysis will focus on the single-user decoding (SUD) scheme where the transmitted signal of each user is decoded separately by the receiver(s) who treat the incoming signal of other users as additive (Gaussian) noise. The main reason for using SUD instead of successive interference cancellation (SIC) is that the former is known to have lower decoding complexity and signalling overhead than the latter – a consequence of SUD not having to broadcast the decoding order to the transmitters [14]. As a result, SIC-based schemes suffer from scalability issues, especially when there are several receivers and/or the channel is highly time-varying.

In this context, non-cooperative power allocation games for static Gaussian interference channels (ICs) have been studied in a series of related papers [7], [8]. There, the existence of a Nash equilibrium is a consequence of the convexity properties of the users’ achievable rates and follows from the general theory of [15]. In fact, under suitable (but stringent) conditions on the channel matrices, it was shown that this equilibrium is unique and that iterative water-filling algorithms converge to it.

Formally, the static PMAC is a special case of this IC frame-

work, but the conditional analysis of [7], [8] almost always fails for static PMAC models. Thus, although the (global) capacity region of this channel is well-understood [4], [5], [2], the channel's distributed version remains unresolved. A first attempt to remedy this was carried out in [11] where it was shown that an associated power allocation game admits an exact potential function [16] whose minimum corresponds to the game's Nash equilibria. However, this potential is *not* strictly convex, so Nash equilibrium uniqueness might fail along with the uniqueness conditions of [7], [8]. Rather surprisingly, it turns out that this is *not* the case: even though these conditions do not hold in the static PMAC context, the Nash equilibrium of the static PMAC game is unique (Theorem 1).

In itself, uniqueness allows us to characterize the system's behavior at equilibrium, but it does not provide a way of actually getting there. Regarding such convergence issues, the authors of [6] considered a single channel with pricing and exhibited power control algorithms which converge to equilibrium under "mild-interference" conditions. Similarly, one of the main results of [12] was to show that if the transmitters know the local channel state and the overall interference-plus-noise covariance matrix, then, subject to similar "mild-interference" conditions, iterative water-filling converges to the equilibrium set of the game (a result which was then enhanced in [17] by dropping this condition for a modified water-filling scheme).

Instead of taking a water-filling approach, we present a simpler learning scheme based on the replicator dynamics of evolutionary game theory [18] which involves the same (often less) information from the side of the players, and which does not require them to solve a nonlinear water-filling problem. Dynamics of this type have been studied extensively in finite [19] and continuous population games [20], but, in *nonlinear games* (such as the one we have here), their properties are not as well understood. Nonetheless, by taking a modified version of the users' utilities, we show that the replicator dynamics converge to an ε -neighborhood of the game's (a.s.) unique equilibrium exponentially fast, i.e. in time $\mathcal{O}(\log(1/\varepsilon))$ (Theorem 3).

In the context of fading however, the static game and the corresponding replicator dynamics lose much of their relevance because variations due to fading open the door to stochasticity. To account for this randomness, we study both fast and block-fading models. Using techniques from the theory of stochastic approximation [21], we show that by properly adjusting their learning scheme, users converge to the (unique) equilibrium of an averaged game whose payoff functions correspond to the users' achievable ergodic rates (Theorems 4 and 5).

Our convergence analysis is based on a novel game-theoretic result which is of independent interest: *in games which admit a (star-)convex potential function, the replicator dynamics converge to (the game's unique) equilibrium at an exponential rate* (Theorem 6). To the best of our knowledge, this is the fastest convergence rate that has been established for the replicator dynamics in the current state of the art [20].

Notational Conventions: If \mathbb{R}^S is the vector space spanned by the set $S = \{s_\alpha\}_{\alpha=1}^S$ and $\{e_\alpha\}_{\alpha=1}^S$ denotes its

canonical basis, we will use α to refer interchangeably to either s_α or e_α , and we will identify the set $\Delta(S)$ of probability measures on S with the standard $(S-1)$ -dimensional simplex of \mathbb{R}^S : $\Delta(S) \equiv \{x \in \mathbb{R}^S : \sum_\alpha x_\alpha = 1 \text{ and } x_\alpha \geq 0\}$. Finally, we will employ Latin indices for players (k, ℓ, \dots) , while reserving Greek ones (α, β, \dots) for their ("pure") strategies; and when summing over $\alpha \in \mathcal{A}_k$, we will simply write $\sum_\alpha^k \equiv \sum_{\alpha \in \mathcal{A}_k}$.

II. SYSTEM MODEL

Following [11], the basic setup of our model is as follows: we consider a finite set $\mathcal{K} = \{1, \dots, K\}$ of wireless single-antenna transmitters (the *players* of the game) who wish to transmit to a group of single-antenna receivers (possibly clustered as a single receiver). Each of these receivers operates on a given channel $\alpha \in \mathcal{A} \equiv \{1, \dots, A\}$ (assumed to be orthogonal, typically in the frequency domain), and each user $k \in \mathcal{K}$ may transmit over a subset $\mathcal{A}_k \subseteq \mathcal{A}$ of these channels (with $A_k \equiv \text{card}(\mathcal{A}_k) \geq 2$).

In particular, if $x_{k\alpha} \sim \mathcal{CN}(0, p_{k\alpha})$ is the transmitted message of user k on channel $\alpha \in \mathcal{A}_k$ and $h_{k\alpha}$ denotes the respective channel coefficient, then the received signal on channel α will be $y_\alpha = \sum_k h_{k\alpha} x_{k\alpha} + z_\alpha$, where $z_\alpha \sim \mathcal{CN}(0, \sigma_\alpha^2)$ denotes the thermal noise. Accordingly, user $k \in \mathcal{K}$ can split his transmitting power among the channels $\alpha \in \mathcal{A}_k$ subject to the constraint:

$$\sum_\alpha^k p_{k\alpha} \leq P_k, \quad (1)$$

where $p_{k\alpha} = \mathbb{E}[|x_{k\alpha}|^2]$ represents the power with which the user transmits on channel α , and P_k is his maximum power. As a result, the *power allocation* of user k will be given by the point $p_k = \sum_\alpha^k p_{k\alpha} e_{k\alpha} \in \mathbb{R}^{\mathcal{A}_k}$, and, analogously, the *power profile* which collects all users' power allocations will be represented by $p = (p_1, \dots, p_K) \in \prod_k \mathbb{R}^{\mathcal{A}_k} = \mathbb{R}^Q$ with $Q \equiv \sum_k A_k$.

In this context, our performance metric will be the users' achievable transmission rates, which depend on their signal to interference-plus-noise ratio (SINR):

$$\text{sinr}_{k\alpha}(p) = \frac{g_{k\alpha} p_{k\alpha}}{\sigma_\alpha^2 + \sum_{\ell \neq k} g_{\ell\alpha} p_{\ell\alpha}}, \quad (2)$$

with $g_{k\alpha} = |h_{k\alpha}|^2$ denoting the channel gain coefficient of user k in channel $\alpha \in \mathcal{A}_k$. Clearly, the users' achievable rates will depend on their power allocation policies through their SINR, but the exact dependence hinges on the time-variability of the channel gain coefficients $g_{k\alpha}$.

At one end, we will study *static* channels, i.e. channels whose coherence time is much larger than both the self-decodable block duration and the power updating period. At the other extreme, we will also consider *fast-fading* channels where the coherence time is much shorter than those characteristic times; here, what matters is the ergodic value of the SINR and the corresponding rate. Finally, we will also analyze the more interesting intermediate case, where the coherence time is greater than the block length but comparable to the update time – hence allowing blocks to be decoded using the instantaneous channel values in (2), but also introducing stochasticity in the game.

A. Static Channels

We will start with the case of static channels employing the single-user decoding (SUD). In this case, the spectral efficiency of user k in the power profile p will be given by [11], [6]:

$$u_k(p) = \sum_{\alpha} b_{\alpha} \log(1 + \text{sinr}_{k\alpha}(p)) \quad (3)$$

where $b_{\alpha} > 0$ represents the bandwidth of channel $\alpha \in \mathcal{A}_k$ and the channel gains $g_{k\alpha}$ are drawn once and for all from a continuous distribution on $[0, \infty)$ at the outset of the game and remain fixed for the duration of the transmission [7], [11]. Then, to maximize their spectral efficiency (3), users saturate (1) by transmitting at the highest possible power [11], so we are led to the *static* PMAC game $\mathfrak{G} \equiv \mathfrak{G}(\mathcal{K}, \{\Delta_k\}, \{u_k\})$ where:

- 1) The *players* of \mathfrak{G} are the transmitters $\mathcal{K} = \{1, \dots, K\}$.
- 2) The *strategy space* of player k is the (scaled) simplex $\Delta_k \equiv P_k \Delta(\mathcal{A}_k) = \{p_k \in \mathbb{R}^{\mathcal{A}_k} : p_{k\alpha} \geq 0 \text{ and } \sum_{\alpha} p_{k\alpha} = P_k\}$ of power allocation vectors; the game's space of strategy profiles $p = (p_1, \dots, p_K)$ will then be $\Delta \equiv \prod_k \Delta_k$.
- 3) The players' *payoffs* (or *utilities*) are given by the spectral efficiencies $u_k: \Delta \rightarrow \mathbb{R}$ of (3).

Of course, the game \mathfrak{G} defined in this way does not adhere to the original normal form of Nash in the sense that *a*) players are not mixing probabilities over a finite set of possible actions, and *b*) even though the players' strategy spaces are simplices, their payoffs are not (multi)linear. On the other hand, with u_k being concave in p_k , it is easy to see that \mathfrak{G} is itself *concave* in the sense of Rosen [15]. Also, as was shown in [11], \mathfrak{G} possesses an *exact potential* [16], i.e. a function $\Phi: \Delta \rightarrow \mathbb{R}$ such that:

$$u_k(p_{-k}; p'_k) - u_k(p_{-k}; p_k) = \Phi(p_{-k}; p_k) - \Phi(p_{-k}; p'_k), \quad (4)$$

for all users $k \in \mathcal{K}$, and for all power allocations $p_k, p'_k \in \Delta_k$ of user k and $p_{-k} \in \Delta_{-k} \equiv \prod_{\ell \neq k} \Delta_{\ell}$ of k 's opponents; in fact, the analysis of [11] also provides the expression:

$$\Phi(p) = - \sum_{\alpha} b_{\alpha} \log(1 + \sum_k g_{k\alpha} p_{k\alpha} / \sigma_{\alpha}^2). \quad (5)$$

B. Fading Channels

As discussed above, the non-static models that we will examine are block-fading and fast-fading channels.

1) *Block-fading channels*: In this case, the coefficients $g_{k\alpha} \equiv g_{k\alpha}(t)$ remain constant over an entire transmission block, so, assuming the transmitter knows (2) for each block through feedback, the users' utilities are still given by (3), with different gains $g_{k\alpha}$ at each self-decodable block.¹ As such, (5) is still a potential for the (now evolving) game $\mathfrak{G}(t)$, the only difference being that Φ will evolve over time following the channels and the game.

2) *Fast-fading channels*: In this regime, the coefficients $h_{k\alpha} \equiv h_{k\alpha}(t)$ evolve ergodically at a rate which is much faster than the characteristic length of a transmission block, so the "instantaneous" utilities (3) lose their relevance. Instead, and assuming for simplicity that users saturate their power

¹Note that we do not assume any delay constraints at the receiver [22]; in this way, reliable information (a la Shannon) can be transmitted over each block.

constraints, their utilities will be given by the *ergodic rates* of [23]:

$$\bar{u}_k(p) = \sum_{\alpha} b_{\alpha} \mathbf{E}_g [\log(1 + \text{sinr}_{k\alpha}(p))]. \quad (6)$$

We thus obtain the *ergodic game* $\bar{\mathfrak{G}} \equiv (\mathcal{K}, \{\Delta_k\}, \{\bar{u}_k\})$, which has the same strategic structure as its static counterpart \mathfrak{G} but payoffs given by (6) instead of (3).

In fact, as in the static case, $\bar{\mathfrak{G}}$ admits the exact potential:

$$\bar{\Phi}(p) \equiv \mathbf{E}_g[\Phi(p)] = - \sum_{\alpha} b_{\alpha} \mathbf{E}_g [\log(1 + \sum_k g_{k\alpha} p_{k\alpha} / \sigma_{\alpha}^2)], \quad (7)$$

whose form depends on the law of the $g_{k\alpha}$. Thus, with $h_{k\alpha} \sim \mathcal{CN}(0, \sqrt{\gamma_{k\alpha}})$, $\gamma_{k\alpha} \geq 0$, the coefficients $g_{k\alpha} = |h_{k\alpha}|^2$ will be χ^2 -distributed, and the calculations of [24, eq. (11)] yield:

Proposition 1. *In i.i.d. Gaussian fast-fading channels with $h_{k\alpha} \sim \mathcal{CN}(0, \sqrt{\gamma_{k\alpha}})$, the ergodic potential $\bar{\Phi}$ is:*

$$\bar{\Phi}(p) = - \sum_{k,\alpha} b_{\alpha} \zeta(r_{k\alpha}^{-1}) \prod_{\ell \neq k} (1 - r_{\ell\alpha} / r_{k\alpha})^{-1} \quad (8)$$

where $r_{k\alpha} = \gamma_{k\alpha} p_{k\alpha} / \sigma_{\alpha}^2$ and $\zeta(x) \equiv \int_0^{\infty} (x+t)^{-1} e^{-t} dt = -e^x \text{Ei}(-x)$.

This proposition will be crucial in the numerical calculations of Section V. For posterity, we only note here that (8) implies that $\bar{\Phi}$ is strictly convex [25], even though, in general, Φ is not.

III. EQUILIBRIUM ANALYSIS

We begin with the notion of *Nash equilibrium*:

Definition 1. A power profile $q \in \Delta$ will be a *Nash equilibrium* of the game \mathfrak{G} (resp. $\bar{\mathfrak{G}}$) when:

$$u_k(q) \geq u_k(q_{-k}; q'_k), \quad (\text{resp. } \bar{u}_k(q) \geq \bar{u}_k(q_{-k}; q'_k)) \quad (9)$$

for all $k \in \mathcal{K}$ and for all $q'_k \in \Delta_k$. In particular, if q satisfies the strict version of (9) for all $q'_k \neq q_k$, it will be called *strict*.

Given that \mathfrak{G} (resp. $\bar{\mathfrak{G}}$) admits a convex potential, its *equilibrium set* will coincide with the minimum set of Φ (resp. $\bar{\Phi}$) [26]. As such, the existence of an equilibrium is guaranteed, and this is already important from a practical point of view because learning protocols would never converge otherwise. Our goal in this section will be to show that these equilibria are essentially *unique*, thus ensuring the system's *predictability* – a crucial feature for performance evaluation, QoS guarantees, etc.

A. Static Channels

With regards to the static potential Φ , it is easy to see that two power profiles $p, p' \in \Delta$ will have $\Phi(p) = \Phi(p')$ whenever

$$\sum_k g_{k\alpha} (p'_{k\alpha} - p_{k\alpha}) = 0 \quad \text{for all } \alpha \in \mathcal{A}. \quad (10)$$

In that case, Φ will not be strictly convex, so its minimum set might fail to be a singleton as well. More precisely, if we set $z = p' - p$, we will also have $\sum_{\alpha} z_{k\alpha} = 0$ for all $k \in \mathcal{K}$, so, on account of (10) above, Φ will not be strictly convex if the following linear system admits a non-zero solution in z :

$$\sum_k g_{k\alpha} z_{k\alpha} = 0, \quad \alpha \in \mathcal{A}; \quad \sum_{\alpha} z_{k\alpha} = 0, \quad k \in \mathcal{K}. \quad (11)$$

Since $z \in \mathbb{R}^Q$, $Q = \sum_k A_k$, and the above $A + K$ constraints are independent (a.s.), we see that if $Q - A - K > 0$, then Φ

cannot be strictly convex – see [27] for more details. In fact, the quantity $\text{ind}(\mathfrak{G}) \equiv Q - A - K$ will be called the *degeneracy index* of the game \mathfrak{G} , and the condition $\text{ind}(\mathfrak{G}) > 0$ means that *if the number of links Q exceeds the number of channels plus transmitters $A + K$, then the game's potential is not strictly convex.*

In typical uplink scenarios of practical interest (e.g. single-receiver OFDM), each user can access all channels, so $A_k = A$ for all k , and, hence $Q = AK > A + K$ (except in small 2×2 systems). This implies that degeneracy appears almost always, so in the absence of strict convexity, a promising way to determine whether the PMAC game admits a *unique* equilibrium would be to use the conditions of [7], [8], [28] where one constructs a certain matrix \mathbf{S} from the channel gain coefficients and tries to show that said matrix has a spectral radius $\rho(\mathbf{S}) < 1$. However, as is shown in [27], this spectral radius exceeds 1 (a.s.), so the results of [7], [8], [28] do not apply to our problem. Still, we have:

Theorem 1. *The static game \mathfrak{G} admits a unique Nash equilibrium (a.s.).*

Sketch of Proof: Let $p \in \Delta$ and consider the (multi)graph $\mathfrak{G}(p) = (\mathcal{A}, \mathcal{E}(p))$ whose vertices are the network's receivers and whose edge set $\mathcal{E}(p)$ is the *multiset sum* $\mathcal{E}(p) = \uplus_k \mathcal{E}_k(p_k)$ where each $\mathcal{E}_k(p_k)$ is a star graph on the nodes $\alpha \in \mathcal{A}$ to which p_k assigns positive power $p_{k\alpha} > 0$ (i.e. all channels to which k transmits with positive power are star-connected and these graphs are superimposed for all $k \in \mathcal{K}$). If p is equilibrial, $\mathfrak{G}(p)$ has to be acyclic [27], so p must lie in the interior of an at most $(A - 1)$ -dimensional face of Δ (a.s.); our assertion then follows from a dimension-counting argument (see [27] for details). ■

B. Fading Channels

1) *Block-fading channels:* As we discussed in Section II-B, the time-varying version of (5) which corresponds to block-fading channels $g_{k\alpha} \equiv g_{k\alpha}(t)$ is a potential for the block-fading game $\mathfrak{G}(t)$. Accordingly, Theorem 1 implies:

Corollary 1. *At each channel realization, the block-fading game $\mathfrak{G}(t)$ admits a unique Nash equilibrium (a.s.).*

2) *Fast-fading channels:* On the other hand, the averaging effect in the ergodic rates (6) can be used to show that the ergodic potential $\bar{\Phi}$ is, in fact, *strictly* convex. This gives:

Theorem 2 ([25]). *The ergodic game $\bar{\mathfrak{G}}$ admits a unique Nash equilibrium.*

IV. LEARNING DYNAMICS AND CONVERGENCE TO EQUILIBRIUM

Although Theorems 1 and 2 guarantee equilibrium uniqueness, it is not at all clear whether users will be able to calculate this equilibrium in decentralized environments where only partial/local information is available at the terminal (e.g., as in distributed or partially distributed cognitive radio networks). Consequently, our goal in this section will be to present a simple distributed learning scheme which allows users to converge to equilibrium, and to determine the speed of this convergence.

This question has attracted considerable interest from the point of view of learning, and two of the most well-studied paradigms are best-response (BR) algorithms and reinforcement learning (RL) [29], [19], [20]. In standard BR schemes [20], players are assumed to monitor their opponents' power allocation policies and respond optimally to them (with respect to their individual utilities). Unfortunately (and in addition to the “perfect monitoring” requirement), it is quite hard to calculate these best responses in large games, so the applicability of this approach to large decentralized networks is quite limited. To circumvent these limitations (in static channels at least), a promising solution lies in the water-filling approach of [12], [7], [8], [17] where users only need their local channel and overall noise-plus-interference covariance matrix. In that case however *a)* users must solve a non-convex fixed point problem at each step; and *b)* convergence is conditional on the interference being low enough (except in [17]). In fact, the conditions of [7], [8] do *not* hold in the PMAC case [27], while the approach of [12] breaks down for large numbers of users [30].

On the other hand, RL algorithms (such as regret-matching [31]) rely on the players knowing their (possibly fictitious) payoffs. Thanks to this information (which, however, is often hard to come by), these algorithms enjoy strong convergence properties in potential games. However, such learning algorithms have been designed for discrete action sets, so it is very hard to adapt them to games with continuous action spaces (such as the ones we are considering here).

To overcome these limitations, our starting point will be the replicator dynamics of evolutionary game theory [18], [19], [20]. The reinforcement aspect of these dynamics does suffer from the same drawback as most RL algorithms (i.e. it applies only to *finite* action sets), but, by exploiting the simplicial structure of the game and its potential, we derive a learning scheme which applies to *continuous* action spaces and which allows users to converge to equilibrium *unconditionally* and *exponentially quickly* (Theorems 3 and 5).

A. Static Channels

Since the replicator equation applies to discrete sets (such as \mathcal{A}_k), a reasonable channel-specific utility would be:

$$u_{k\alpha}(p) = b_\alpha \log(1 + \text{sinr}_{k\alpha}(p)) \quad (12)$$

which leads to the replicator equation:

$$\frac{dp_{k\alpha}}{dt} = p_{k\alpha}(t) \left(u_{k\alpha}(p(t)) - P_k^{-1} \sum_{\beta} p_{k\beta}(t) u_{k\beta}(p(t)) \right), \quad (13)$$

whose second term ensures that $p(t) \in \Delta$ for all $t \geq 0$. Unfortunately, the utility u_k of eq. (3) is *not* a convex combination of the $u_{k\alpha}$, so (13) is not well-behaved w.r.t. the game \mathfrak{G} either – for instance, Nash equilibria are not stationary.

Instead, given that each user invariably seeks to unilaterally increase his utility, we will consider the *marginal utilities*:

$$v_{k\alpha}(p) \equiv \frac{\partial u_k}{\partial p_{k\alpha}} = \frac{b_\alpha g_{k\alpha}}{\sigma_\alpha^2 + \sum_{\ell} g_{\ell\alpha} p_{\ell\alpha}}. \quad (14)$$

Since player k can calculate $v_{k\alpha}(p)$ by means of $\text{sinr}_{k\alpha}$ and $g_{k\alpha}$ alone (the bandwidths b_α are assumed fixed and known),

any dynamics based on the $v_{k\alpha}$'s will be inherently distributed (and simpler than solving a water-filling problem to boot). We will thus consider the replicator equation:

$$\frac{dp_{k\alpha}}{dt} = p_{k\alpha}(t)(v_{k\alpha}(p(t)) - v_k(p(t))), \quad (15)$$

where v_k denotes the user average $v_k(p) = P_k^{-1} \sum_{\beta} p_{k\beta} v_{k\beta}(p)$.

Remark 1 (Comparison to other power updating schemes). The replicator equation (15) is clearly quite unlike the water-filling schemes of [12], [7], [8], [17]. Closer in spirit to (15) are the algorithms developed in the 90's with the goal of minimizing transmitting power by comparing the instantaneous SINR to a target value and iteratively updating the power proportionally to this difference [32], [33], [34]; still, there is little overlap with these algorithms, both in terms of setup and convergence.

Remark 2. The marginal utilities $v_{k\alpha}$ are similar but *not* equal to the SINR (2) of user k at a given channel $\alpha \in \mathcal{A}_k$, and they do not coincide with the popular metric of total rate per unit power either [13], [3]. These metrics can all be calculated based on the same feedback (and might appear more appealing than $v_{k\alpha}$), but we shall see that it is precisely the (perhaps unconventional) choice of the marginal utilities $v_{k\alpha}$ that leads to convergence.

The first important property of (15) is that its rest points satisfy the (waterfilling) condition $v_{k\alpha}(p) = v_{k\beta}(p)$ for all nodes $\alpha, \beta \in \text{supp}(p_k)$ to which user k transmits with positive power. Hence, by the Karush-Kuhn-Tucker (KKT) conditions of [11], we see that *Nash equilibria of \mathfrak{G} are stationary in (15)*.

The converse of this statement is not true: every vertex of Δ is stationary *without* necessarily being a Nash equilibrium. Nonetheless, the game's (unique) Nash equilibrium is the *only* attracting state of the dynamics (see Appendix A for the proof):

Theorem 3. *Let $q \in \Delta$ be the (a.s.) unique equilibrium of \mathfrak{G} . Then, every interior solution orbit of the replicator dynamics (15) converges to q ; moreover, there exists $c > 0$ such that*

$$D_{\text{KL}}(q \| p(t)) \leq D_{\text{KL}}(q \| p(0)) e^{-ct} \text{ for all } t \geq 0, \quad (16)$$

where D_{KL} is the Kullback-Leibler divergence. In other words, replicator trajectories converge to an ε -neighborhood of a Nash equilibrium in time $\mathcal{O}(\log(1/\varepsilon))$.

Remark 1. To the best of our knowledge, the exponential convergence rate of Theorem 3 (see also Theorem 6 in App. A and Fig. 1) is the fastest known estimate for the replicator dynamics (see [20] for a review of the state of the art).

Remark 2. The Kullback-Leibler divergence is defined as [20]:

$$D_{\text{KL}}(q \| p) = \sum_{k, \alpha: q_{k\alpha} > 0} q_{k\alpha} \log(q_{k\alpha}/p_{k\alpha}). \quad (17)$$

Clearly, $D_{\text{KL}}(q_k \| p_k)$ is finite if and only if p_k allocates positive power $p_{k\alpha} > 0$ to all channels $\alpha \in \text{supp}(q)$ which are present in q_k . Thus, in particular, Theorem 3 guarantees that uniform initial power allocations equilibrate exponentially quickly.

Remark 3. In the evolutionary analysis of [35], the fitness of a species (the number of descendants in the unit of time) might be nonlinear, but it is still a convex combination of each

phenotype's fitness. In this special case, the dynamics of [35] are formally equivalent to (15), and it is shown therein that their limit points are Nash equilibria. Theorem 3 (see also Theorem 6) extends this analysis by demonstrating that the replicator dynamics really *do* converge to Nash equilibrium, and that the rate of this convergence is exponential.

Remark 4. It was shown in [7] that iterative water-filling algorithms converge exponentially when the water-filling operator is a contraction. However, given that the sufficient conditions which guarantee the contraction property fail in the PMAC case [27], the analysis of [7] does not apply here.

Of course, the value of the exponent of (16) is critical because it controls how fast users converge to equilibrium. Thus, if we consider the "instantaneous" convergence exponents:

$$\lambda_k(t) \equiv -\frac{1}{t} \log \frac{D_{\text{KL}}(q_k \| p_k(t))}{D_{\text{KL}}(q_k \| p_k(0))}, \quad (18)$$

then Theorem 3 simply states that the total equilibration rate $\lambda(t) \equiv \min_k \{\lambda_k(t)\}$ is at least c . For the sake of simplicity, we will only present here an analytic expression for the value of c for *strict* equilibria (for the full analysis including non-strict equilibria, see Appendix A). In this case, if user k transmits with full power to channel α_k at equilibrium, we will have:

$$c_k \equiv \liminf_t \{\lambda_k(t)\} = \gamma_k^{-1} (1 - e^{-\gamma_k}) \Delta v_k \text{ and } c = \min_k \{c_k\} \quad (19)$$

where $\Delta v_k = \min \{v_{k, \alpha_k}(q) - v_{k\beta}(q) : \beta \neq \alpha_k\}$ is k 's minimum deviation cost and $\gamma_k = D_{\text{KL}}(q \| p(0))/P_k$. We thus obtain:

Proposition 2. *If $q = \sum_k P_k e_{k, \alpha_k}$ is a strict equilibrium of the static game \mathfrak{G} , the power of user k on channel α_k grows as:*

$$p_{k, \alpha_k}(t) \sim P_k (1 - e^{-\Delta v_k t}). \quad (20)$$

Proof: From (19), we have $c_k \rightarrow \Delta v_k$ as $\gamma_k \rightarrow 0$. However, since $D_{\text{KL}}(q_k \| p_k(t)) \rightarrow 0$ as $t \rightarrow \infty$ (Theorem 3), we will have $\gamma_k \rightarrow 0$ by definition, and our assertion follows. ■

B. Fading Channels

1) *Block-fading channels:* In this case, the replicator equation (15) becomes non-deterministic because the coefficients $g_{k\alpha}$ evolve stochastically over time. To account for this, we will rewrite the replicator dynamics (15) in discrete time as:

$$\Delta p_{k\alpha}(n+1) = \delta(n) p_{k\alpha}(n) [v_{k\alpha}(p(n), g(n)) - v_k(p(n), g(n))], \quad (21)$$

where $\Delta p_{k\alpha}(n+1) \equiv p_{k\alpha}(n+1) - p_{k\alpha}(n)$ and the "step" $\delta(n)$ is a (possibly time-dependent) learning parameter.

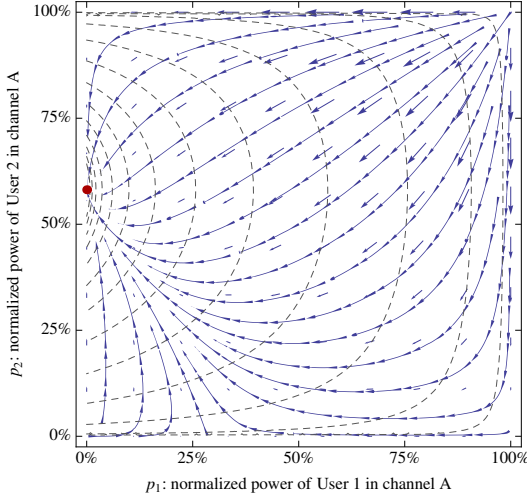
For simplicity, we will concentrate here on the case where the temporal variations of the channels are uncorrelated. In this case, if we set $\bar{v}_{k\alpha} = \mathbf{E}[v_{k\alpha}]$ and $\eta_{k\alpha} = v_{k\alpha} - \bar{v}_{k\alpha}$, we obtain:

$$\begin{aligned} p_{k\alpha}(n+1) &= p_{k\alpha}(n) + \delta(n) p_{k\alpha}(n) (\bar{v}_{k\alpha}(p(n)) - \bar{v}_k(p(n))) \\ &\quad + \delta(n) p_{k\alpha}(n) (\eta_{k\alpha}(p(n), g(n)) - \eta_k(p(n), g(n))), \end{aligned} \quad (22)$$

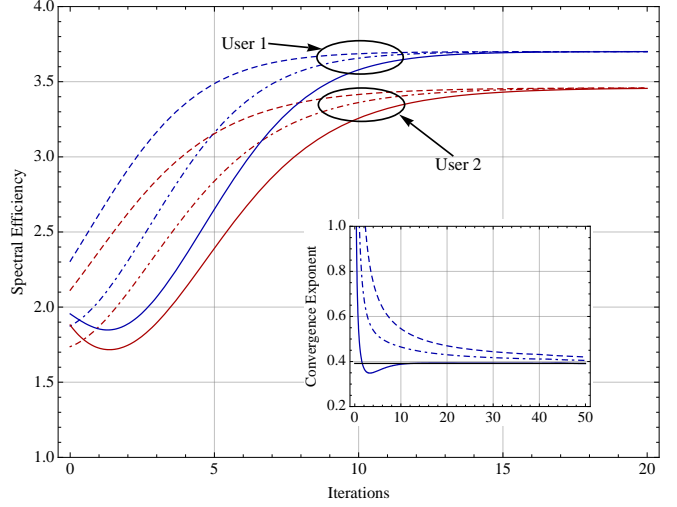
where the $\bar{v}_{k\alpha}$ are deterministic and the $\eta_{k\alpha}$ are zero-mean. In fact, if we interchange expectation and differentiation, we get:

$$\bar{v}_{k\alpha}(p) = \mathbf{E} \left[\frac{\partial u_k}{\partial p_{k\alpha}} \right] = \frac{\partial}{\partial p_{k\alpha}} \mathbf{E}[u_k] = \frac{\partial \bar{u}_k}{\partial p_{k\alpha}}, \quad (23)$$

Convergence of the replicator dynamics in a 2x2 system



(a) Convergence of the replicator dynamics in static channels.



(b) Spectral efficiency over time for different initial configurations.

Fig. 1. Convergence to equilibrium in a 2×2 game with static channels. The dashed contours in Fig. 1(a) are the level sets of the K-L divergence w.r.t. the game's equilibrium (red dot), while p_1, p_2 represent the normalized power allocation of each user. In Fig. 1(b), we plot the spectral efficiency u_k of each user as a function of time for three randomly drawn initial configurations; as can be seen in the inlay, the equilibration rate $\lambda(t)$ coincides with the value predicted by Theorem 3 (solid black line).

so the mean utilities of (22) are the gradients of the ergodic rates (6). Thus, if we remove the noise $\eta_{k\alpha}$ from (22), the general theory of [21] shows that (22) will track the *mean-field equation*:

$$\frac{dp_{k\alpha}}{dt} = p_{k\alpha}(t) [\bar{v}_{k\alpha}(p(t)) - \bar{v}_k(p(t))], \quad (24)$$

so the asymptotic properties of (22) will follow those of (24).

We thus see that the asymptotic behavior of the replicator algorithm (21) for block-fading channels is intimately linked to the ergodic game $\bar{\mathcal{G}}$. In itself, this is quite natural because the ergodic equilibrium of $\bar{\mathcal{G}}$ represents the only reasonable time-invariant equilibrated notion for the block-fading game $\mathcal{G}(t)$ with temporally uncorrelated channels (see also [36]). More to the point, we have (see App. B for the proof):

Theorem 4. *If the learning parameters $\delta(n)$ of (21) satisfy $\sum_{n=1}^{\infty} \delta(n) = \infty$ and $\sum_{n=1}^{\infty} \delta^2(n) < \infty$, then the algorithm (21) for temporally uncorrelated block-fading channels converges (a.s) to the (unique) Nash equilibrium of the ergodic game $\bar{\mathcal{G}}$.*

Remark. The most usual choice for the parameters $\delta(n)$ is $\delta(n) = 1/n$. These variable rates can be interpreted either as the actual time step of the algorithm, or as a discount that users apply to their updating scheme at every tick of a timer. This last interpretation is crucial for practical purposes because there are hard limits to how fast a device can update its policy. For constant $\delta(n) \equiv \delta$, the dynamics (21) evolve faster, but convergence to equilibrium is in the distribution sense of [21].

2) *Fast-fading channels:* As we have already noted in Section II-B, the users' (ergodic) rates (6) in the fast-fading regime depend on the channels' statistics, so instantaneous channel information obtained when updating their powers is of little use. Because of this, the system becomes effectively deterministic, so, similarly to the static case, users may base their learning on the (mean) replicator learning dynamics (24) – i.e. the discrete-time learning scheme (21) without the noise

term and with a constant step $\delta(n) = \delta$. We thus obtain (see App. A):

Theorem 5. *The mean dynamics (24) converge to the (unique) Nash equilibrium of the ergodic game $\bar{\mathcal{G}}$, and this convergence is exponential: interior orbits ε -equilibrate in time $\mathcal{O}(\log(1/\varepsilon))$.*

Remark. If the ergodic equilibrium is strict, the convergence exponent (18) of the mean dynamics (24) is just (App. A):

$$\bar{c} = \min_k \{\bar{c}_k\} \quad \text{with} \quad \bar{c}_k = \gamma_k^{-1} (1 - e^{-\gamma_k}) \Delta \bar{v}_k, \quad (25)$$

where $\gamma_k = D_{\text{KL}}(q \| p(0))/P_k$ as in (19), but now $\Delta \bar{v}_k = \min \{\bar{v}_{k,\alpha_k}(q) - \bar{v}_{k\beta}(q) : \beta \neq \alpha_k\}$ is the minimum deviation cost of user k for the *mean* marginal utilities $\bar{v}_{k\alpha}$.

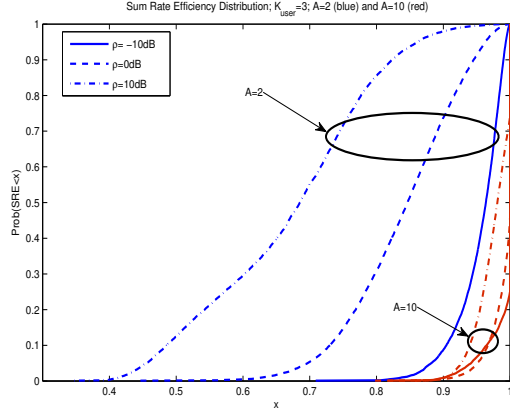
V. NUMERICAL SIMULATIONS

In this section, our aim is to validate our theoretical results by means of numerical simulations. We begin by introducing the *sum-rate efficiency* (SRE) of a power profile p :

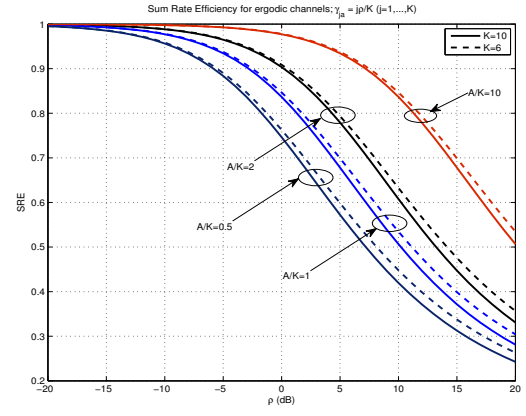
$$\text{SRE}(p) = \frac{\sum_k u_k(p)}{C_{\text{sum}}}, \quad (\text{resp. } \overline{\text{SRE}}(p) = \frac{\sum_k \bar{u}_k(p)}{\bar{C}_{\text{sum}}} \text{ for } \bar{\mathcal{G}}) \quad (26)$$

i.e. the ratio of the sum of achievable rates in the power profile p over the maximum achievable aggregate sum-rate under SIC (which is the sum-capacity of the multiple access channel (MAC)); interestingly, if q is the game's (unique) equilibrium, then $C_{\text{sum}} = -\Phi(q)$ (and similarly for the ergodic case).

In Fig. 2(a), we plot the SRE at Nash equilibrium for randomly drawn static channels. While the equilibrium SRE can deviate significantly from its maximum value (unity) for $A < K$, in the $A \geq K$ regime, the SRE is typically close to 100% (and, in fact, equal to 100% with positive probability). We may attribute this to the fact that for $A \geq K$ there is a finite probability that the system's equilibrium is at a vertex of Δ where each user is alone on a single channel, inducing

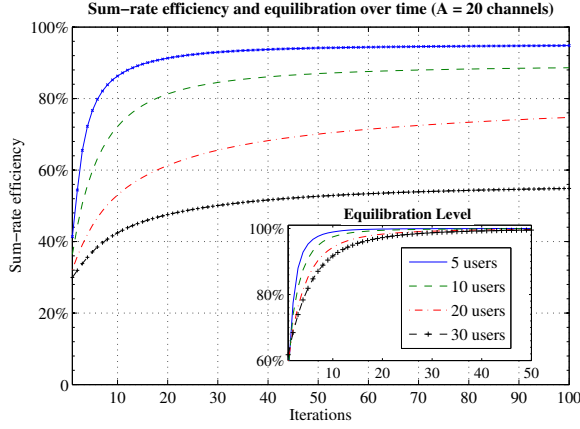


(a) Equilibrium sum-rate efficiency for static channels.

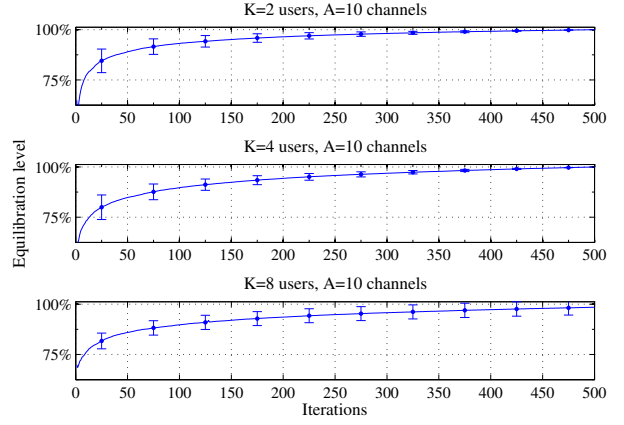


(b) Equilibrium sum-rate efficiency for ergodic channels.

Fig. 2. The CDF of the equilibrium sum-rate efficiency for static channels (Fig. 2(a)) and the equilibrium SRE for ergodic Gaussian i.i.d. channels as a function of the thermal SNR parameter $\rho = P_{\max}/\sigma^2$ (Fig. 2(b)) for different numbers of channels A and users K (all with similar maximum power constraints).



(a) Sum-rate efficiency and equilibration over time for static channels.



(b) Long-term equilibration in block-fading channels.

Fig. 3. Equilibration (and its efficiency) for different numbers of users and channels. In static channels, the replicator dynamics equilibrate extremely fast, even for a large number of users; in temporally uncorrelated block-fading channels, the dynamics still converge, but slower – due to the discounting $\delta(n) = 1/n$ in (21).

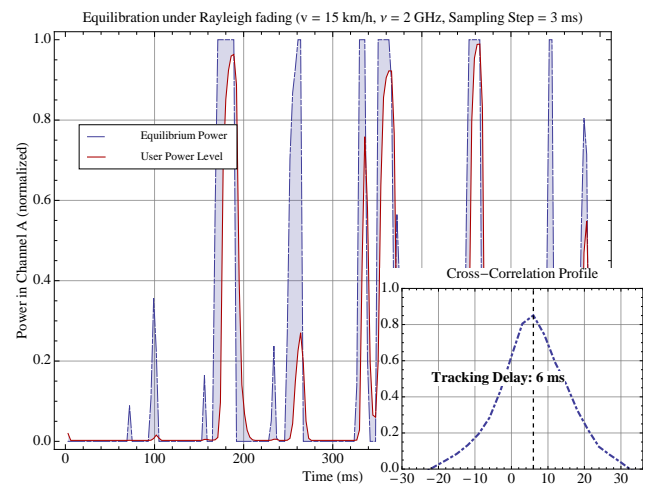
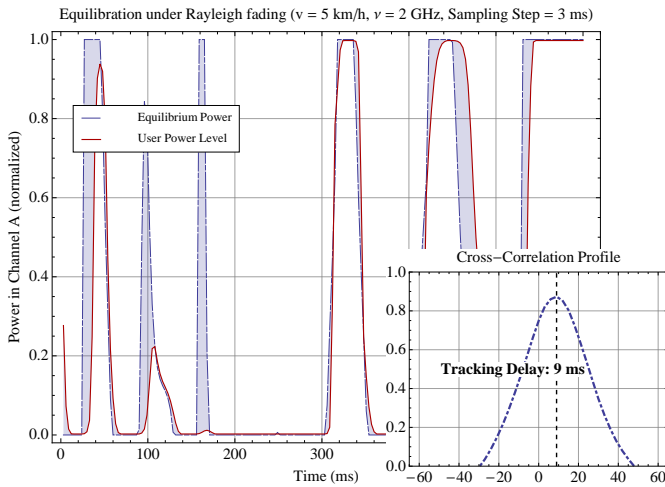


Fig. 4. The Nash power level (dashed blue) and the actual power level learned by a user (solid red) in a 2×2 Jakes-fading game. As can be seen by the cross-correlograms of the two time-series (inlays), users are away from equilibrium for only a very short amount of time (light blue shading): we observe a 9 ms tracking delay for user velocities in the 5 km/h range and 6 ms for 15 km/h, meaning that the replicator dynamics converge within 10-15% of the system's coherence time (108 ms and 36 ms respectively).

optimal performance. Thus, with fair probability (close to 1/2 based on our simulations) the complex SIC scheme yields no performance benefit over the much simpler SUD approach. Analogously, in Fig. 2(b), we plot the SRE at equilibrium for ergodic channels by using Proposition 1 to evaluate the maximum sum-rate under SIC; in that case, while the SRE is nearly optimal for small SNR, it deviates strongly from its maximum value for larger SNR values.

Furthermore, to test the equilibration rate of (21) for different values of K and A , we introduce the equilibration level (EQL):

$$\text{EQL}(p) = \Phi(p)/\Phi(q), \text{ (resp. } \overline{\text{EQL}}(p) = \overline{\Phi}(p)/\overline{\Phi}(q) \text{ for } \overline{\mathfrak{G}}), \quad (27)$$

with q being the equilibrium of \mathfrak{G} (resp. $\overline{\mathfrak{G}}$) – so an EQL of 1 implies that the system has reached its equilibrium.

Beginning with the static case, in Fig. 3(a) we drew $N = 50$ channel realizations and ran the discrete-time learning scheme (21) with constant $g(n)$ and $\delta(n)$ for $A = 20$ channels and $K = 5, 10, 20$ and 30 users. Then, by plotting the average SRE and EQL over time, we see that even for 30 users, the system equilibrates within a few tens of iterations. On the other hand, for the ergodic block-fading scenario of Fig. 3(b), we used $A = 10$ channels and plotted the system's $\overline{\text{EQL}}$ over time for $K = 2, 4$ and 8 users learning with $\delta(n) = 1/n$. As predicted by Theorem 4, the system converges to the game's ergodic equilibrium – but slower due to the discounting $\delta(n) = 1/n$. Finally, for fast-fading users following the discrete-time version of (24), the EQL and SRE plots were virtually identical to the static case (to be expected since both dynamical systems are deterministic), so they have been omitted for space considerations.

Finally, to test the convergence rate of the replicator dynamics in more realistic (non-ergodic) fading conditions that do not possess a long-term stationary equilibrium, we also simulated in Fig. 4 channels that follow the well-known Jakes model for Rayleigh fading [37]. Specifically, we considered a 2×2 game with user velocities $v = 5$ km/h and 15 km/h (Figs. 4(a) and 4(b) respectively) transmitting at a carrier frequency of $\nu = 2$ GHz. We then ran the learning scheme (21) with a constant update period of $\delta = 3$ ms, and we plotted the (normalized) power level of a single user against the evolving equilibrium power level (calculated at each step based on the instantaneous channel coefficients). Then, to quantify how well the users follow the system's evolving equilibrium, we calculated the cross-correlation of the two processes and the users' *tracking delay* (defined as the point of maximum cross-correlation).

Remarkably, Fig. 4 shows that the dynamics track the game's evolving equilibrium extremely closely. On average, users equilibrate within 9 ms for 5 km/h fading velocities, and within 6 ms for 15 km/h – meaning that users converge within 10-15% of the system's coherence time (108 ms and 36 ms respectively).

Remark. For larger numbers of users (or channels), the results observed are similar; for instance, if the users in the previous game are increased to a few tens (we went up to $N = 50$), their tracking delay becomes longer but never exceeds 30-35% of

the channel coherence time. However, due to space limitations, we opted to present here only the 2×2 case for simplicity.

VI. CONCLUSIONS AND FUTURE DIRECTIONS

In this paper, we studied the distributed power allocation problem for orthogonal uplink channels by introducing a game which admits a convex potential function. For both static and fading channels, we found that the associated game admits a unique Nash equilibrium and we showed that a simple distributed learning scheme based on the replicator dynamics converges to equilibrium from (almost) any initial condition. In fact, by proving a general result for convex potential games, we showed that the speed of this convergence is exponential: users converge to an ε -neighborhood of an equilibrium in time which is at most of order $\mathcal{O}(\log(1/\varepsilon))$.

There is a number of important extensions of this work which demonstrate the strength of the replicator dynamics in continuous nonlinear games of this sort. First off, instead of the achievable rates u_k , one could consider energy-efficient metrics where users do not saturate their power constraints – e.g., when the price of transmission power might restrain users from transmitting at maximum power. More importantly, these techniques can be extended even to non-orthogonal channel models such as the multiple-input multiple-output (MIMO) MAC case where the game's strategy space consists of all positive-definite precoding matrices with constrained trace. Because of this nonlinear structure, the form (15) of the replicator dynamics no longer applies, but one can still write down a suitably modified *matrix-valued* replicator equation which allows users to converge to equilibrium.

APPENDIX A

CONVERGENCE SPEED OF THE REPLICATOR DYNAMICS

Recall first that the *solid tangent cone* to Δ at q is the set of rays starting at q and intersecting Δ in at least one other point, i.e.: $T_q^c \Delta \equiv \{z \in \mathbb{R}^Q : z_{k\alpha} \geq 0 \text{ for all } \alpha \in \mathcal{A}_k \text{ with } q_{k\alpha} = 0\}$. With this in mind, we have the following generalization of convexity:

Definition 2. A function $F: \Delta \rightarrow \mathbb{R}$ will be called *star-convex* w.r.t. $q \in \Delta$ if $f(\theta) \equiv F(q + \theta z)$ is convex and increasing for all $z \in T_q^c \Delta$ and for all $\theta > 0$ s.t. $q + \theta z \in \Delta$.

Star-convex functions need not be convex, but strictly convex functions are star-convex w.r.t. their global minimum and weakly convex functions with a unique minimum are also star-convex – in particular, both $\Phi, \overline{\Phi}$ are star-convex. For games with star-convex potentials, we then have:

Theorem 6. Let $\mathfrak{Q} \equiv \mathfrak{Q}(\mathcal{K}, \{\Delta_k\}, \{\phi_k\})$ be a game with a star-convex potential F . Then, the replicator dynamics (15) for the marginal utilities $\phi_{k\alpha} = \frac{\partial \phi_k}{\partial p_{k\alpha}}$ converge to q for any initial condition that starts at finite K -L divergence $h_0 \equiv D_{\text{KL}}(q \| p(0))$ from q . Moreover, there exists $c > 0$ such that:

$$D_{\text{KL}}(q \| p(t)) \leq h_0 e^{-ct} \text{ for all } t \geq 0. \quad (28)$$

Our proof strategy will be to establish an inequality of the form $\frac{d}{dt} H_q(p(t)) \leq -c H_q(p(t))$ and then employ Grönwall's

lemma. To that end, we first define the “evolutionary index”:

$$L_q(p) = -\sum_{k,\alpha} (p_{k\alpha} - q_{k\alpha}) \phi_{k\alpha}(p), \quad (29)$$

so named because q is evolutionarily stable iff $L_q(p) > 0$ near q . In fact, if we set $H_q(p) = D_{\text{KL}}(q \| p)$, an easy calculation shows that L_q is just (the negative of) the time derivative of H_q w.r.t. the replicator dynamics (15): $\frac{d}{dt} H_q(p(t)) = \sum_{k,\alpha} \frac{\partial H_q}{\partial p_{k\alpha}} \dot{p}_{k\alpha} = -L_q(p(t))$.

We will therefore begin by showing that $L_q(p) > 0$ for all $p \in \Delta \setminus \{q\}$, implying that H_q is Lyapunov for the replicator dynamics (15), and proving the convergence part of Theorem 6. Indeed, if we fix some $z \in T_q^c \Delta$, (29) may be rewritten as $f'(\theta) = \sum_{k,\alpha} \frac{\partial F}{\partial p_{k\alpha}} \Big|_{q+\theta z} z_{k\alpha} = \theta^{-1} L_q(q+\theta z)$, for all $\theta > 0$ such that $q+\theta z \in \Delta$. Then, with $f(\theta)$ convex and increasing (Definition 2), we obtain the estimate $L_q(p) = \theta f'(\theta) \geq f(\theta) - f(0) = F(p) - F(q)$, which shows that $L_q(p) > 0$ for all $p \neq q$.

To prove the convergence time estimate (28) we will need to show that L_q grows linearly along directions which are not supported in q , and quadratically along those which are supported in q . To be specific, let $V_q = \{x \in \mathbb{R}^Q : x_{k\alpha} = 0 \text{ if } q_{k\alpha} = 0\}$ be the subspace of directions of \mathbb{R}^Q , $Q = \sum_k A_k$, which are supported in q , and let V_q^\perp be its orthocomplement in \mathbb{R}^Q . Then, by decomposing $z \in \mathbb{R}^Q$ as $z = z_\parallel + z_\perp$ with $z_\parallel \in V_q$ and $z_\perp \in V_q^\perp$, we define the seminorms $\|\cdot\|_\parallel$ and $\|\cdot\|_\perp$ as:

$$\|z\|_\parallel^2 \equiv \|z_\parallel\|_2^2 = \sum_{k,\alpha} z_{k\alpha}^2, \quad \|z\|_\perp \equiv \|z_\perp\|_1 = \sum_{k,\alpha} |z_{k\alpha}|, \quad (30)$$

where the notation $\sum_{k,\alpha}^\parallel, \sum_{k,\alpha}^\perp$ is shorthand for summing over the directions of V_q and V_q^\perp respectively. We thus get:

Lemma 1. *Let $F: \Delta \rightarrow \mathbb{R}$ be star-convex w.r.t. $q \in \Delta$. Then:*

$$L_q(p) \geq F(p) - F(q) \geq m \|p - q\|_\perp + \frac{1}{2} r \|p - q\|_\parallel^2, \quad (31)$$

where $m = \min_k \{\phi_{k\alpha}(q) - \phi_{k\mu}(q) : q_{k\mu} = 0, q_{k\alpha} > 0\}$, and r is the minimum of the Rayleigh quotient $\langle z, \mathbf{M}(q+z)z \rangle / \|z\|^2$ for the Hessian $\mathbf{M}(p) = \frac{\partial^2 F}{\partial p_{k\alpha} \partial p_{l\beta}}$ of F , restricted over $T_q^c \Delta$.

Proof: Since q minimizes F , the KKT conditions give $\phi_{k\alpha}(q) = -\frac{\partial F}{\partial p_{k\alpha}} \Big|_q = -\lambda_k$ for all $\alpha \in A_k$ such that $q_{k\alpha} > 0$ and $\phi_{k\alpha}(q) < -\lambda_k$ otherwise (where λ_k denotes the complementary slackness Lagrange multiplier of F over Δ). Thus, a first order Taylor estimate with Lagrange remainder readily yields:

$$f(\theta) = f(0) + f'(0)\theta + \frac{1}{2} f''(\xi)\theta^2 \quad (32)$$

for some $\xi \in (0, \theta)$, so (31) will follow once we properly estimate the linear and quadratic terms of (32).

As far as the linear term of (32) is concerned, we will have $f'(0) = \sum_{k,\alpha} z_{k\alpha} \frac{\partial F}{\partial p_{k\alpha}} \Big|_q = \sum_{k,\alpha}^\parallel z_{k\alpha} \frac{\partial F}{\partial p_{k\alpha}} \Big|_q + \sum_{k,\alpha}^\perp z_{k\alpha} \frac{\partial F}{\partial p_{k\alpha}} \Big|_q = \sum_{k,\alpha}^\perp z_{k\alpha} \left(\frac{\partial F}{\partial p_{k\alpha}} \Big|_q - \lambda_k \right) \geq m \|z\|_\perp$, where the last equality holds because $\sum_{k,\alpha}^\perp z_{k\alpha} = -\sum_{k,\alpha}^\parallel z_{k\alpha}$ (recall that $z \in T_q^c \Delta$) and the last inequality is just the definition of m . Similarly, for any $\xi \in (0, \theta)$ and $z \in T_q^c \Delta$, we get $f''(\xi) = \langle z, \mathbf{M}(q+\xi z)z \rangle = R_{q+\xi z}(\xi z) \|z\|^2$, where $R_p(w) = \langle w, \mathbf{M}(p)w \rangle$, $p \in \Delta$, $w \in T_p \Delta$, denotes the Rayleigh quotient of the Hessian \mathbf{M} of F . Hence, if r is the minimum of $R_{q+w}(w)$ over the set $B_q = \{w \in T_q^c \Delta : q+w \in \Delta\}$, we will also have $f''(\xi) \geq r \|z\|^2$, and (31) follows by plugging the above into (32) and noting that $\|z\| \geq \|z\|_\parallel$. ■

Obtaining similar estimates for the relative entropy function H_q is harder (after all, H_q blows up near the boundary of Δ), so we will need two more auxiliary lemmas:

Lemma 2. *For all $z \in T_q^c \Delta \setminus \{0\}$ and for all $a > 1$, the equation*

$$H_q(q+\theta z) = a \|z\|_\perp \theta + \frac{1}{2} a \sum_{k,\beta}^\parallel z_{k\beta}^2 / q_{k\beta} \theta^2, \quad (33)$$

admits a unique positive root $\theta_a \equiv \theta_a(z)$. Consequently:

$$H_q(q+\theta z) \leq a \|z\|_\perp \theta + \frac{1}{2} a \sum_{k,\beta}^\parallel z_{k\beta}^2 / q_{k\beta} \theta^2 \text{ for all } \theta \leq \theta_a(z). \quad (34)$$

Proof: Let $h(\theta) \equiv H_q(q+\theta z)$ be the LHS of (33), and denote its RHS by $ag(\theta)$. Then, if we set $w(\theta) = h(\theta) - ag(\theta)$, we readily obtain $w(0) = 0$, $w'(0) = \|z\|_\perp(1-a) \leq 0$, and $w''(0) = \sum_{k,\beta}^\parallel z_{k\beta}^2 / q_{k\beta}(1-a) < 0$, and the result follows by simple arguments relying on the mean value theorem. ■

Lemma 3. *Let $F: \Delta \rightarrow \mathbb{R}$ be star-convex w.r.t. $q \in \Delta$ and let $p(t)$ be a solution orbit of the replicator dynamics with initial relative entropy $h_0 = H_q(p(0))$. Then, there exists $b > 1$ s.t.:*

$$H_q(p(t)) \leq b \|p(t) - q\|_\perp + \frac{b}{2q_0} \|p(t) - q\|_\parallel^2, \quad (35)$$

where $q_0 = \min_{k,\alpha} \{q_{k\alpha} : q_{k\alpha} > 0\}$.

Proof: Fix some $a > 1$. Then, by Lemma 2, we know that (33) admits a unique positive root $\theta_a(z)$, so let $h_a(z) = H_q(q+\theta_a(z)z)$ and set $h_a = \max\{h_a(z) : z \in S_q\}$, where $S_q = \{z \in T_q^c \Delta : z+q \in \Delta \text{ but } q+(1+\varepsilon)z \notin \Delta \text{ for any } \varepsilon > 0\}$. Moreover, set $h_c = \max\{h_0, h_a\}$, let $\theta_c(z)$ be the unique positive root of the equation $H_q(q+\theta_c(z)z) = h_c$, and define $b(z) = g(\theta_c(z))/h_c$ with $g(\theta) = \|z\|_\perp \theta + \frac{1}{2} \sum_{k,\beta}^\parallel z_{k\beta}^2 / q_{k\beta} \theta^2$ (as in the proof of Lemma 2). We will then have $b(z) \geq a$ since, otherwise, (34) would yield the contradiction $h_c = b(z)g(\theta_c(z)) < ag(\theta_c(z)) < h(\theta_c(z)) = h_c$.

With $b(z) > 1$, a second application of Lemma 2 yields $H_q(q+\theta z) \leq b(z) (\|z\|_\perp \theta + \frac{1}{2} \sum_{k,\beta}^\parallel z_{k\beta}^2 / q_{k\beta} \theta^2)$ for all $\theta \leq \theta_c(z)$. Thus, if we decompose $p(t)$ as $p(t) = q + \theta(t)z(t)$ with $\theta > 0$ and $z(t) \in S_q$, we will have $\theta(t) \leq \theta_c(z(t))$; indeed, should this ever fail, we would have $H_q(p(t)) > b(z(t))g(\theta(t)) > b(z(t))g(\theta_c(t)) = h_c \geq h_0$ which contradicts the fact that H_q is Lyapunov. Hence, with $\theta(t) \leq \theta_c(z(t))$ for all $t \geq 0$, we get $H_q(p(t)) \leq b(z(t)) (\|z(t)\|_\perp \theta(t) + \frac{1}{2} \sum_{k,\beta}^\parallel z_{k\beta}^2(t) / q_{k\beta} \theta^2(t))$, and (35) follows by taking $b = \max\{b(z) : z \in S_q\}$. ■

Proof of Theorem 6: With notation as in Lemmas 1 and 3, let $c = \min\{m/b, r q_0/b\}$. We then get $L_q(p(t)) \geq m \|p(t) - q\|_\perp + \frac{1}{2} r \|p(t) - q\|_\parallel^2 \geq c H_q(p(t))$ and Grönwall's lemma yields $H_q(p(t)) \leq h_0 e^{-ct}$. Since the KKT inequalities for F are strict along any direction of \mathbb{R}^Q which is not supported in q , we will have $m > 0$ and, consequently, $c > 0$ as well. ■

Proof of Theorems 3 and 5: The potentials Φ and $\bar{\Phi}$ are star-convex, so both theorems follow from Theorem 6. ■

All that remains is to calculate the value of c when q is strict. In that case, given that the intersection of V_q with $T_q^c \Delta$ is trivial, the quadratic term of (31) can be ignored and we get $L_q(p) \geq \frac{1}{2} \sum_k \|p_k - q_k\|_1 \Delta \phi_k$, where $\Delta \phi_k = \min_{\mu \neq \alpha_k} \{\phi_{k,\alpha_k}(q) - \phi_{k,\mu}(q)\} > 0$. As for (35), we may decompose $p_k \in \Delta_k \setminus \{q_k\}$ as $p_k = q_k + \theta_k z_k$ where $z_k \in T_{q_k}^c \Delta_k$ has $z_{k,\alpha_k} = -P_k$. Thus, with $p_{k,\alpha_k} = P_k(1 - \theta_k)$, we readily obtain $H_q(p) = -\sum_k P_k \log(1 - \theta_k)$. Now, let θ_k^* be defined by the equation $h_0 = H_{q_k}(q_k + \theta_k z_k)$, i.e., $\theta_k^* = 1 - \exp(-h_0/P_k)$, implying that $-P_k \log(1 - \theta_k) \leq$

$h_0\theta_k/\theta_k^*$ iff $0 \leq \theta_k \leq \theta_k^*$ (because of convexity). We then claim that $H_q(p(t)) = -\sum_k P_k \log(1 - \theta_k(t)) \leq h_0 \sum_k \theta_k(t)/\theta_k^*$, where $\theta_k(t)$ is defined via the decomposition $p_k(t) = q_k + \theta_k(t)z_k(t)$.

However, if $\theta_k(t) > \theta_k^*$ for some $t \geq 0$, then we would have $H_q(p_k(t)) > h_0$, and, hence, $H_q(p(t)) > H_q(p(0))$ as well, a contradiction – recall that $H_q(p(t))$ is decreasing. Thus, combining all the above, we only need pick c such that $P_k \Delta\phi_k \geq ch_0/\theta_k^*$, and the sharpest such choice is:

$$c = \min_k \left\{ P_k/h_0 \left(1 - e^{-h_0/P_k} \right) \Delta\phi_k \right\}. \quad (36)$$

APPENDIX B

STOCHASTIC APPROXIMATION OF THE REPLICATOR DYNAMICS

Proof of Theorem 4: Note first that Δ is invariant under the dynamics (22) if the $\delta(n)$ are chosen small enough. To see this, we will restrict ourselves w.l.o.g. to a game with one user and two choices, A and B (the general argument being similar). Thus, if we let $p_A(n) \equiv p_{1,A}(n)$ be the power that the user sends to channel A at the n -th iteration of the dynamics, we must find $\delta(n)$ such that $0 \leq p(n) \leq 1$ for all $n \geq 0$ and all possible $g_{A,B}(n) \geq 0$. So, assuming this holds for some $n \geq 0$, we get:

$$p_A(n+1) - p_A(n) = \delta(n) p_A(n)(1 - p_A(n)) \times \left(\frac{g_A(n)}{\sigma_A^2 + g_A(n)p_A(n)} - \frac{g_B(n)}{\sigma_B^2 + g_B(n)(1 - p_A(n))} \right). \quad (37)$$

The first term of the LHS of (37) is positive and the second is uniformly bounded, say by M , so $\delta(n) \leq M$ yields $p_A(n+1) \geq 0$. The complementary inequality $p_A(n+1) \leq 1$ then follows similarly, so, with $p(n) \in \Delta$ for all n , our theorem follows from Theorem 2 and Corollary 4 in Chap. 2 of [21]. ■

REFERENCES

- [1] S. Lasaulce, M. Debbah, and E. Altman, "Methodologies for analyzing equilibria in wireless games," *IEEE Signal Process. Mag.*, vol. 26, no. 5, pp. 41–52, September 2009.
- [2] G. A. Gupta and S. Toupis, "Power allocation over parallel Gaussian multiple access and broadcast channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3274–3283, July 2006.
- [3] F. Meshkati, H. V. Poor, and S. C. Schwartz, "Energy-efficient resource allocation in wireless networks: an overview of game-theoretic approaches," *IEEE Signal Process. Mag.*, vol. Special Issue on Resource-Constrained Signal Processing, Communications and Networking, May 2007.
- [4] R. S. Cheng and S. Verdú, "Gaussian multiaccess channels with ISI: capacity region and multiuser water-filling," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 773–785, May 1993.
- [5] D. N. Tse and S. V. Hanly, "Multiaccess fading channels Part I: polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inf. Theory*, vol. 44, no. 11, pp. 2796–2815, November 1998.
- [6] T. Alpcan, T. Başar, R. Srikant, and E. Altman, "CDMA uplink power control as a noncooperative game," *Wireless Networks*, vol. 8, pp. 659–670, 2002.
- [7] G. Scutari, D. P. Palomar, and S. Barbarossa, "Competitive design of multiuser MIMO systems based on game theory: a unified view," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1089–1103, September 2008.
- [8] —, "The MIMO iterative waterfilling algorithm," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1917–1935, May 2009.
- [9] S. Chiochan, E. Hossain, and J. Diamond, "Channel assignment structures for infrastructure-based 802.11 WLANs: a survey," *IEEE Commun. Surveys Tuts.*, vol. 12, no. 1, p. 124, first quarter 2010.
- [10] S. Lasaulce, A. Suarez, R. de Lacerda, and M. Debbah, "Using cross-system diversity in heterogeneous networks: throughput optimization," *Elsevier Journal of Performance Evaluation*, vol. 65, no. 11, pp. 907–921, December 2008.
- [11] S. M. Perlaza, E. V. Belmega, S. Lasaulce, and M. Debbah, "On the base station selection and base station sharing in self-configuring networks," in *ValueTools '09: Proceedings of the 4th International Conference on Performance Evaluation Methodologies and Tools*, 2009.
- [12] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 5, pp. 1105–1115, May 2002.
- [13] F. Meshkati, M. Chiang, H. V. Poor, and S. C. Schwartz, "A game-theoretic approach to energy-efficient power control in multicarrier CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 24, pp. 1115–1129, 2006.
- [14] E. V. Belmega, "On resource allocation games in MIMO wireless channels," Ph.D. dissertation, Université Paris Sud 11, 2010.
- [15] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave n -person games," *Econometrica*, vol. 33, no. 3, pp. 520–534, 1965.
- [16] D. Monderer and L. S. Shapley, "Potential games," *Games and Economic Behavior*, vol. 14, no. 1, pp. 124–143, 1996.
- [17] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–152, 2004.
- [18] P. D. Taylor and L. B. Jonker, "Evolutionary stable strategies and game dynamics," *Mathematical Biosciences*, vol. 40, no. 1–2, pp. 145–156, 1978.
- [19] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*, ser. Economic learning and social evolution. Cambridge, MA: The MIT Press, 1998, vol. 2.
- [20] W. H. Sandholm, *Population Games and Evolutionary Dynamics*, ser. Economic learning and social evolution. Cambridge, MA: MIT Press, 2011.
- [21] V. S. Borkar, *Stochastic approximation*. Cambridge University Press and Hindustan Book Agency, 2008.
- [22] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 10, pp. 359–378, May 1994.
- [23] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986–1992, 1997.
- [24] A. L. Moustakas and S. H. Simon, "Optimizing Multiple-Input Single-Output (MISO) communication systems with general Gaussian channels: nontrivial covariance and nonzero mean," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2770–2780, October 2003.
- [25] E. V. Belmega, S. Lasaulce, M. Debbah, and A. Hjørungnes, "Distributed power allocation policies in MIMO channels," in *EUSIPCO '10: Proceedings of the 2010 European Signal Processing Conference*, 2010.
- [26] A. Neyman, "Correlated equilibrium and potential games," *International Journal of Game Theory*, vol. 26, pp. 223–227, 1997.
- [27] P. Mertikopoulos, E. V. Belmega, A. L. Moustakas, and S. Lasaulce, "Dynamic power allocation games in parallel multiple access channels," in *ValueTools '11: ACM Proceedings of the 5th International Conference on Performance Evaluation Methodologies and Tools*, 2011.
- [28] Z.-Q. Luo and J.-S. Pang, "Analysis of iterative waterfilling algorithms for multi-user power control in digital subscriber lines," *EURASIP J. Appl. Signal Process. [Online]*, May 2006.
- [29] T. Börgers and R. Sarin, "Learning through reinforcement and replicator dynamics," *Journal of Economic Theory*, vol. 77, pp. 1–14, 1997.
- [30] J. Huang and Z. Han, *Cognitive Radio Networks: Architectures, Protocols, and Standards*. Auerbach Publications, CRC Press, 2010, ch. Game theory for spectrum sharing.
- [31] S. Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, vol. 68, no. 5, pp. 1127–1150, September 2000.
- [32] G. J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 641–646, November 1993.
- [33] N. Bambos, S. C. Chen, and G. J. Pottie, "Channel access algorithms with active link protection for wireless communication networks with power control," *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp. 583–57, October 2000.
- [34] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 13, pp. 1341–1347, September 1995.
- [35] W. H. Sandholm, "Potential games with continuous player sets," *Journal of Economic Theory*, vol. 97, pp. 81–108, 2001.
- [36] P. Mertikopoulos and A. L. Moustakas, "The emergence of rational behavior in the presence of stochastic perturbations," *The Annals of Applied Probability*, vol. 20, no. 4, pp. 1359–1388, 2010.

- [37] G. Calcev, D. Chizhik, B. Göransson, S. Howard, H. Huang, A. Kogiantis, A. F. Molisch, A. L. Moustakas, D. Reed, and H. Xu, "A wideband spatial channel model for system-wide simulations," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, p. 389, March 2007.



Panayotis Mertikopoulos received the B.S. degree in physics with a major in astrophysics from the University of Athens in 2003, the M.Sc. and M.Phil. degrees in mathematics from Brown University in 2005 and 2006 respectively, and the Ph.D. degree in physics from the University of Athens in 2010. In 2010-2011 he was with the Economics Department of École Polytechnique. He is currently a CNRS Researcher at the Laboratoire d'Informatique de Grenoble, France.

Dr. Mertikopoulos has been an Embeirikeion Foundation Fellow since 2003. His research interests lie in dynamical systems, stochastic analysis, game theory, and their applications to networks.



Elena V. Belmega received the B.S. degree from the University Politehnica of Bucharest, România, in 2007. She obtained the M.Sc. and Ph.D. degrees, both from the Université Paris-Sud 11, Paris, France, in 2007 and 2010 respectively. In 2010-2011, Dr. Belmega was a post-doctoral researcher in a joint project between the Alcatel-Lucent Chair on Flexible Radio in Supélec and Princeton University. She is currently an assistant professor at ETIS/ENSEA – Université de Cergy-Pontoise – CNRS, Cergy-Pontoise, France.

Dr. Belmega was one of the recipients of the 2009 L'Oréal – UNESCO – French Academy of Sciences Fellowship: "For young women doctoral candidates in science".



Aris Moustakas (SM'04) received the B.S. degree in physics from Caltech in 1990 and the M.S. and Ph.D. degrees in theoretical condensed matter physics from Harvard University in 1992 and 1996, respectively. He joined Bell Labs, Lucent Technologies, NJ, in 1998, first in the Physical Sciences Division and then also in the Wireless Advanced Technology Laboratory. He is currently an assistant professor at the Physics Department of the University of Athens, Greece.

Dr. Moustakas is serving as Associate Editor for the *IEEE TRANSACTIONS ON INFORMATION THEORY*. His main research interests lie in the areas of multiple antenna systems and the applications of game theory and statistical physics to the theory of communications and networks.



Samson Lasaulce received the B.S. degree in Applied Physics from École Normale Supérieure (Cachan) and the M.Sc. and Ph.D. degrees in Signal Processing from École Nationale Supérieure des Télécommunications, Paris, France. He was with Motorola Labs from 1999 to 2001 and with France Télécom R&D in 2002 and 2003. Since 2004, he is with CNRS and Supélec as a Senior Researcher and École Polytechnique as professor.

Dr. Lasaulce is the recipient of the best student paper awards in *VALUETOOLS '07* and *CROWNCOM '09*, and the best paper award of *NETGCOOP '11*. He organized *GAMECOMM '09* and *WNC3 '08*, and was the general chair of *VALUETOOLS '11*. Dr. Lasaulce is serving as Associate Editor for the *IEEE TRANSACTIONS ON SIGNAL PROCESSING*. His research interests lie in communications, signal processing and information theory with a focus on game theory for wireless communications.